CLAIMS

What is claimed is:

1. A method for performing database recovery after a crash of an instance of a database, wherein multiple transactions were active when the instance crashed, the method comprising the steps of:

identifying a plurality of dead transactions;

determining statistical data about said plurality of dead transactions;

determining that a particular number of recovery servers should be used to recover said plurality of dead transactions based on the statistical data; and

recovering said plurality of dead transactions using said particular number of recovery servers.

2. The method of Claim 1, wherein the step of recovering said plurality of dead transactions is performed by executing the particular number of recovery servers in parallel.

3. The method of Claim 1, wherein:

the step of identifying the plurality of dead transactions includes the step of maintaining a working list, wherein the working list identifies a list of dead transactions for which recovery will be attempted; and

the step of determining statistical data includes the step of determining statistical data based on the list of dead transactions.

4. The method of Claim 3, wherein the step of maintaining a working list comprises the steps of:

locating a rollback segment, wherein the rollback segment contains a transaction table that contains entries associated with dead transactions;

scanning the transaction table to identify the dead transactions; and

storing the identity of the dead transactions in the working list.

5. The method of Claim 4, wherein:

the method further comprises the step of maintaining a block count, wherein the block count identifies the number of undo blocks that are associated with a particular transaction; and

the step of determining statistical data includes the step of determining a total number of undo blocks that need to be recovered, wherein the total number of undo blocks is based on the block count associated with the dead transactions identified in the working list.

6. The method of Claim 4, wherein the step of determining statistical data includes the step of determining statistical data based on the number of dead transactions that are identified in the working list.

7. The method of Claim 1, wherein the step of determining that the particular number of recovery servers should be used includes the step of determining that the particular number of recovery servers should be used based on a max_parallelism threshold value, wherein the max_parallelism threshold value provides an upper limit for the number of recovery servers to be used.

8. The method of Claim 7, further comprises the step of determining the max_parallelism threshold value based on a user input value.

9. The method of Claim 1, further comprising the steps of:

identifying a rollback segment that was previously owned by the crashed instance at the time of its crash; and

the crashed instance reacquiring ownership of the rollback segment after the crashed instance is restarted.

10. The method of Claim 9, wherein the step of reacquiring ownership of the rollback segment includes the steps of:

identifying an instance that currently owns the rollback segment that was previously owned by the crashed instance at the time of its crash;

requesting the instance to release ownership of the rollback segment; and

the instance releasing ownership of the rollback segment in response to the request.

11. The method of Claim 1, further comprises the steps of:

identifying a rollback segment that is unowned, wherein the unowned rollback segment is not currently associated with any instance of the database; and

associating the unowned rollback segment with the crashed instance, wherein associating the unowned rollback segment with the crashed node causes the rollback segment to be owned by the crashed instance.

12. The method of Claim 1, wherein the step of recovering the plurality of dead transactions comprises the steps of:

maintaining a working list, wherein the working list identifies a list of dead transactions for which recovery will be attempted;

selecting a dead transaction from the working list;

acquiring a rollback segment lock on a rollback segment, wherein the rollback segment is associated with a transaction table that contains an entry that corresponds to the dead transaction;

acquiring a transaction lock on a chain of undo, wherein the chain of undo contain change information associated with the dead transaction;

determining whether the dead transaction still needs to be recovered; and

if the dead transaction still needs to be recovered, assigning the dead transaction to a recovery server.

13. The method of Claim 12, wherein the step of acquiring the transaction lock includes the step of a coordinator acquiring the transaction lock.

14. The method of Claim 13, wherein the method further comprises the steps of:

upon completing the recovery of the dead transaction, the recovery server signaling the coordinator to indicate it has completed the recovery of the dead transaction; and

upon receiving the signal from the recovery server, the coordinator releasing its lock on the transaction.

15. The method of Claim 1, wherein the step of recovering the plurality of dead transactions using the particular number of recovery servers includes the steps of:

assigning two or more dead transactions to a recovery server;

associating a time slice value with the recovery server, wherein the time slice value is used by the recovery server to promote fairness during recovery of the two or more dead transactions; and

recovering the two or more dead transactions using the time slice value.

16. A computer-readable medium carrying one or more sequences of one or more instructions for performing database recovery after a crash of an instance of a database, wherein multiple transactions were active when the instance crashed, wherein the execution of the one or more sequences of one or more instructions by one or more processors causes the one or more processors to perform the steps of:

identifying a plurality of dead transactions;

determining statistical data about said plurality of dead transactions;

determining that a particular number of recovery servers should be used to recover said plurality of dead transactions based on the statistical data; and

recovering said plurality of dead transactions using said particular number of recovery servers.

17. The computer-readable medium of Claim 16, wherein the step of recovering said plurality of dead transactions is performed by executing the particular number of recovery servers in parallel.

18. The computer-readable medium of Claim 16, wherein:

the step of identifying the plurality of dead transactions includes the step of maintaining a working list, wherein the working list identifies a list of dead transactions for which recovery will be attempted; and

the step of determining statistical data includes the step of determining statistical data based on the list of dead transactions.

19. The computer-readable medium of Claim 18, wherein the step of maintaining a working list comprises the steps of:

locating a rollback segment, wherein the rollback segment contains a transaction table that contains entries associated with dead transactions;

scanning the transaction table to identify the dead transactions; and

storing the identity of the dead transactions in the working list.

20. The computer-readable medium of Claim 19, wherein:

the computer-readable medium further comprises instructions for performing the step of maintaining a block count, wherein the block count identifies the number of undo blocks that are associated with a particular transaction; and

the step of determining statistical data includes the step of determining a total number of undo blocks that need to be recovered, wherein the total number of undo blocks is based on the block count associated with the dead transactions identified in the working list.

21. The computer-readable medium of Claim 19, wherein the step of determining statistical data includes the step of determining statistical data based on the number of dead transactions that are identified in the working list.

22. The computer-readable medium of Claim 16, wherein the step of determining that the particular number of recovery servers should be used includes the step of determining that the particular number of recovery servers should be used based on a max_parallelism threshold value, wherein the max_parallelism threshold value provides an upper limit for the number of recovery servers to be used.

23. The computer-readable medium of Claim 22, further comprises instructions for performing the step of determining the max_parallelism threshold value based on a user input value.

24. The computer-readable medium of Claim 16, further comprising instructions for performing the steps of:

identifying a rollback segment that was previously owned by the crashed instance at the time of its crash; and

the crashed instance reacquiring ownership of the rollback segment after the crashed instance is restarted.

25. The computer-readable medium of Claim 24, wherein the step of reacquiring ownership of the rollback segment includes the steps of:

identifying an instance that currently owns the rollback segment that was previously owned by the crashed instance at the time of its crash;

requesting the instance to release ownership of the rollback segment; and

the instance releasing ownership of the rollback segment in response to the request.

26. The computer-readable medium of Claim 16, further comprises instructions for performing the steps of:

identifying a rollback segment that is unowned, wherein the unowned rollback segment is not currently associated with any instance of the database; and

associating the unowned rollback segment with the crashed instance, wherein associating the unowned rollback segment with the crashed node causes the rollback segment to be owned by the crashed instance.

27. The computer-readable medium of Claim 16, wherein the step of recovering the plurality of dead transactions comprises the steps of:

maintaining a working list, wherein the working list identifies a list of dead transactions for which recovery will be attempted;
selecting a dead transaction from the working list;
acquiring a rollback segment lock on a rollback segment, wherein the rollback segment is associated with a transaction table that contains an entry that corresponds to the dead transaction;
acquiring a transaction lock on a chain of undo, wherein the chain of undo contain change information associated with the dead transaction;
determining whether the dead transaction still needs to be recovered; and
if the dead transaction still needs to be recovered, assigning the dead transaction to a recovery server.

28. The computer-readable medium of Claim 27, wherein the step of acquiring the transaction lock includes the step of a coordinator acquiring the transaction lock.

29. The computer-readable medium of Claim 28, wherein the computer-readable medium further comprises instructions for performing the steps of:

upon completing the recovery of the dead transaction, the recovery server signaling the coordinator to indicate it has completed the recovery of the dead transaction; and
upon receiving the signal from the recovery server, the coordinator releasing its lock on the transaction.

30. The computer-readable medium of Claim 16, wherein the step of recovering the plurality of dead transactions using the particular number of recovery servers includes the steps of:

assigning two or more dead transactions to a recovery server;

associating a time slice value with the recovery server, wherein the time slice value is used by the recovery server to promote fairness during recovery of the two or more dead transactions; and

recovering the two or more dead transactions using the time slice value.

31. A system for performing database recovery after a crash of an instance of a database, wherein multiple transactions were active when the instance crashed, the system comprising:

a memory;

one or more processors coupled to the memory; and

a set of computer instructions contained in the memory, the set of computer instructions including computer instructions which when executed by one or more processors, cause the one or more processors to perform the steps of:

identifying a plurality of dead transactions;

determining statistical data about said plurality of dead transactions;

determining that a particular number of recovery servers should be used to recover said plurality of dead transactions based on the statistical data; and

recovering said plurality of dead transactions using said particular number of recovery servers.

32. The system of Claim 31, wherein the step of recovering said plurality of dead transactions is performed by executing the particular number of recovery servers in parallel.

33. The system of Claim 31, wherein:

the step of identifying the plurality of dead transactions includes the step of maintaining a working list, wherein the working list identifies a list of dead transactions for which recovery will be attempted; and

the step of determining statistical data includes the step of determining statistical data based on the list of dead transactions.

34. The system of Claim 33, wherein the step of maintaining a working list comprises the steps of:

locating a rollback segment, wherein the rollback segment contains a transaction table that contains entries associated with dead transactions;

scanning the transaction table to identify the dead transactions; and

storing the identity of the dead transactions in the working list.

35. The system of Claim 34, wherein:

the system further comprises the step of maintaining a block count, wherein the block count identifies the number of undo blocks that are associated with a particular transaction; and

the step of determining statistical data includes the step of determining a total number of undo blocks that need to be recovered, wherein the total number of undo blocks is based on the block count associated with the dead transactions identified in the working list.

36. The system of Claim 31, wherein the step of determining that the particular number of recovery servers should be used includes the step of determining that the particular number of recovery servers should be used based on a max_parallelism threshold value, wherein the max_parallelism threshold value provides an upper limit for the number of recovery servers to be used.

37. The system of Claim 31, further comprising the steps of:

identifying a rollback segment that was previously owned by the crashed instance at the time of its crash; and

the crashed instance reacquiring ownership of the rollback segment after the crashed instance is restarted.

38. The system of Claim 31, further comprises the steps of:

identifying a rollback segment that is unowned, wherein the unowned rollback segment is not currently associated with any instance of the database; and

associating the unowned rollback segment with the crashed instance, wherein associating the unowned rollback segment with the crashed node causes the rollback segment to be owned by the crashed instance.

39. The system of Claim 31, wherein the step of recovering the plurality of dead transactions comprises the steps of:

maintaining a working list, wherein the working list identifies a list of dead transactions for which recovery will be attempted;

selecting a dead transaction from the working list;

acquiring a rollback segment lock on a rollback segment, wherein the rollback segment is associated with a transaction table that contains an entry that corresponds to the dead transaction;

acquiring a transaction lock on a chain of undo, wherein the chain of undo contain change information associated with the dead transaction;

determining whether the dead transaction still needs to be recovered; and

if the dead transaction still needs to be recovered, assigning the dead transaction to a recovery server.

40. The system of Claim 31, wherein the step of recovering the plurality of dead transactions using the particular number of recovery servers includes the steps of:

assigning two or more dead transactions to a recovery server;

associating a time slice value with the recovery server, wherein the time slice value is used by the recovery server to promote fairness during recovery of the two or more dead transactions; and

recovering the two or more dead transactions using the time slice value.